DOCUMENT RESUME

ED 390 397 IR 055 744

AUTHOR Poggione, Marcia Dae
TITLE Retrieving Articles in Series from Medline and SciSearch.
PUB DATE Apr 95
NOTE 73p.; Masters Research Paper, Kent State University.
PUB TYPE Dissertations/Theses - Undetermined (040)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Indexing; *Information Retrieval; Information Systems; *Medical Research; Online Catalogs; Online Searching; Periodicals; *Scholarly Journals; Search Strategies; *Subject Index Terms
IDENTIFIERS Journal Articles; MEDLINE; SCISEARCH; *Special Issues (Periodicals); *Titles

ABSTRACT

Although articles in series generally embrace a central theme, they may not share a common index term; retrieval by medical subject heading often produces an incomplete set. Since subject retrieval is problematic, it would be helpful if articles in a series could be retrieved by series title. Five quality-assessed general medical journals published in 1993 were examined revealing 134 series. The retrieval effectiveness of series titles was measured utilizing formulas for recall and precision in the Medline and SciSearch databases. Mean recall of articles in series from both databases was 23% and mean precision was 25%. Although the mean difference between Medline and SciSearch recall was .07, the slight positive difference exhibited by Medline was not statistically significant. Correlation coefficients demonstrate a positive rather than inverse relationship between recall and precision for both databases. Mean recall of series with attached titles (a title-subtitle configuration) was 94% and unattached titles (journal section title-article title configuration) was 11%. These results demonstrate both databases adhere to documentation pertaining to articles in series. Three editions of "Guide to Special Issues and Indexes of Periodicals" were examined for background information on the indexing of articles in series. Nine figures and eight tables present information. (Contains 61 references.) (AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 390 397

RETRIEVING ARTICLES IN SERIES FROM MEDLINE AND SCISEARCH

A Master's Research Paper submitted to the
Kent State University School of Library Science
in partial fulfillment of the requirements
for the degree Master of Library Science

by

Marcia Dae Poggione

April, 1995

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R. DuMont
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR055744

BEST COPY AVAILABLE

ABSTRACT

Although articles in series generally embrace a central theme, they may not share a common index term. Retrieval by medical subject heading often produces an incomplete set. Since subject retrieval is problematic, it would be helpful if articles in series could be retrieved by series title. Five quality-assessed general medical journals published in 1993 were examined revealing 134 series. The retrieval effectiveness of series titles was measured utilizing formulas for recall and precision in the Medline and SciSearch databases. Mean recall of articles in series from both databases was 23 percent and mean precision was 25 percent. Although the mean difference between Medline and SciSearch recall was .07, the slight positive difference exhibited by Medline was not statistically significant. Correlation coefficients demonstrate a positive rather than inverse relationship between recall and precision for both databases. Mean recall of series with attached titles (a title-subtitle configuration) was 94 percent and unattached titles (journal section title-article title configuration) was 11 percent. These results demonstrate both databases adhere to documentation pertaining to articles in series. The possible reasons for database documentation noncompliance were discussed. Three editions of *Guide to Special Issues and Indexes of Periodicals* were examined for background information on the indexing of articles in series.

Master's Research Paper by

Marcia Dae Poggione

B.S., Albion College, 1966

M.S., Western Reserve University, 1968

M.L.S., Kent State University, 1995

Approved by

Adviser Thomas J. Frosbach Date April 20, 1995

TABLE OF CONTENTS

LIST OF FIGURES.....v

LIST OF TABLES.....vi

CHAPTER

I. INTRODUCTION.....1

Statement of the Problem.....1

Justification.....5

Statement of Purpose.....6

Limitations.....7

II. LITERATURE REVIEW.....9

Background Information.....9

Indexing the Journal Literature.....11

Indexing Journal Titles.....18

Indexing Article Titles.....21

Indexing Articles in Series.....25

Indexing Inconsistencies.....29

Retrieval Difficulties.....31

III. METHODOLOGY.....36

IV. ANALYSIS OF DATA.....39

General Results.....39

Recall Results.....39

Precision Results.....43

Correlation Between Recall and Precision....44

V. DISCUSSION AND CONCLUSIONS.....45

Retrieval Effectiveness.....45

Recall and Precision Relationship.....48

Human vs Automated Indexing.....49

Database Documentation.....50

Recommendations.....51

APPENDIX

A. Articles in Series Appearing in Figure 1.....54

B. Series Excluded from Study.....57

WORKS CITED.....58

LIST OF FIGURES

Figure

1. Retrieval of a Series and Individual Articles in the Series Using the Series Title in the Medline Database.....4

2. Series Title Attached.....4

3. Series Title Unattached.....4

4. Hooper's Formula.....31

5. Recall Formula.....34

6. Precision Formula.....34

7. Periodicals Included in the Study.....36

8. Article Information.....37

9. Series Information.....37

LIST OF TABLES

Table

1. Percent of Eight Articles within the Series, "Users' Guides Assigned Selected Medical Subject Headings.....2

2. Percentage of Total Series by Series Type and Journal.....40

3. Percentage of Total Series by Series Type and Title Type..40

4. Mean Percent Recall and Precision of Articles in Series by Title Type.....41

5. Mean Percent Recall and Precision of Articles in Series by Series Type.....41

6. Percentage of Series with Unattached Titles Retrieved by Percent Recall.....42

7. Series with Unattached Titles and 50-99 Percent Recall....47

8. Percentage of Series with 0 or 100 Percent Recall and Precision.....49

CHAPTER I

INTRODUCTION

Statement of the Problem

Medical periodicals frequently publish articles in series. They may appear in a special issue, supplement, or section of the periodical. Usually there is a collective title for the entire series as well as individual titles for each article within the series.

In the medical literature, articles in series generally embrace a central theme. However, they may not share a common index term, or this term may not be readily apparent. Thus, retrieval by subject heading often produces an incomplete set. Table 1 illustrates the medical subject headings assigned to a series of eight articles with the collective title, "Users' Guides to the Medical Literature." In this representative example, the only common indexed term is 'periodicals.' Furthermore, only 21 percent of the medical subject headings were assigned to 50 percent or more articles in the series.

Since subject retrieval is problematic, it would be helpful if articles in series could be retrieved by their series title. This would guarantee a complete set. A parallel situation exists with monographs in series. When shelving these items, a decision is made either to keep the series together as a cohesive unit or to place each item in the series under its corresponding subject area. Only

Table 1.--Percent of Eight Articles Within the Series, "Users' Guides"* Assigned Selected Medical Subject Headings

Medical Subject Heading	Number	Percent
Clinical Competence	1	12.5
Clinical Trials	1	12.5
Decision Support Techniques	4	50
Dementia	1	12.5
Diagnosis, Differential	1	12.5
Diagnosis, Laboratory	1	12.5
Diagnostic Services	1	12.5
Education, Medical, Continuing	2	25
Grateful Med	3	37.5
Guidelines	6	75
Information Storage & Retrieval	2	25
Likelihood Functions	2	25
Medical Informatics Applications	6	75
MEDLARS	1	12.5
MEDLINE	3	37.5
Outcome Assessment	1	12.5
Patient Care Planning	4	50
Periodicals	8	100
Physician's Practice Patterns	1	12.5
Prognosis	1	12.5
Reproducibility of Results	6	75
Research	1	12.5
Research Design	1	12.5
Review Literature	1	12.5
Sensitivity and Specificity	1	12.5
Subject Headings	1	12.5
Technology, Medical	3	37.5
Treatment Outcome	2	25

*See appendix A for a list of articles appearing in this series.

one physical location can exist. However, if individual items in the series have been shelved separately, the entire series can be retrieved using the series title added entry. The second edition of *Anglo-American Cataloguing Rules* suggests the generous assignment of a series title added entry:

> Make an added entry under the heading for a series for each separately catalogued work in the series if it provides a useful collocation. . . .In case of doubt, make a series added entry (Gorman and Winkler 1988).

The *Handbook for AACR2* further explains:

> LC [Library of Congress] traces series if such series were published before the twentieth century, if they are entered under personal author, or if they are published by a noncommercial publisher, particularly a small or "alternative" press (Maxwell 1989).

Although similar in some respects, the journal literature presents unique problems. Journal articles are confined to their respective journal issues in much the same way that monographs are confined to a specific shelf location. Unlike monographs, however, journal articles are separated from other articles pertaining to the same subject. Moreover, individual articles within a series are separated from one another, as they frequently appear in different issues of a journal.

Both monographs in series and articles in series have collective titles. However, often there is not a comparable "series title added entry" for a series appearing in the journal literature. Therefore, consistent retrieval by series title is not always possible (see figure 1).

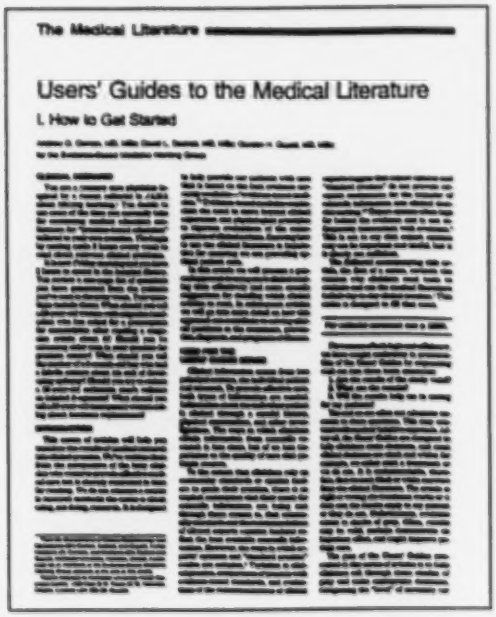
The journal literature does not cooperate with standard formats for a series title. Sometimes the series title appears as

an integral part of the individual article title (see figure 2) and other times it is physically separated (see figure 3).

Figure 1.--Retrieval of a Series and Individual Articles in the Series Using the Series Title in the Medline Database

Series Title*	Type of Series	Retrieval	
		Series	Articles
Users' guides to the medical literature	Limited sequence; title attached	Yes	Yes
Facts, figures, & fallacies	Limited sequence; title unattached	No	No
Outpatient Parenteral Antibiotic Therapy	Supplement	Yes	No
From the Centers of Disease Control & Prevention	Permanent section	Yes	Yes

*For a complete listing of individual articles in each series see appendix A.


The Medical Literature

Users' Guides to the Medical Literature

I. How to Get Started

Figure 2.--Series Title Attached

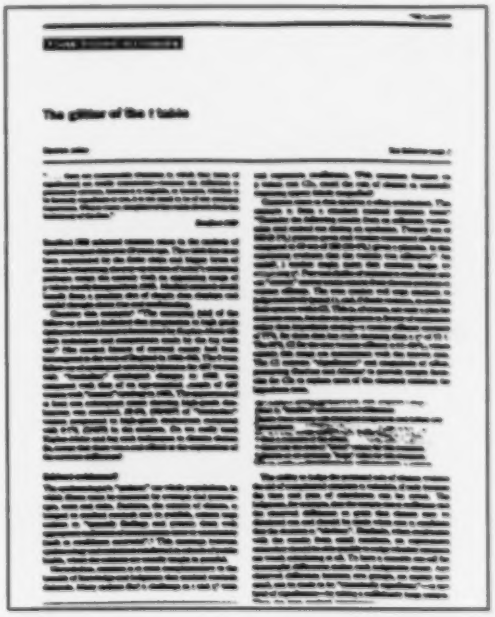


Figure 3.--Series Title Unattached

BEST COPY AVAILABLE

Justification

Experienced medical librarians agree that health care professionals frequently request articles grouped together in series (Augustine 1994; Rosenthal 1994). Disciplines other than medicine also request this type of specialized information. In response to this need, the Special Libraries Association published three editions of *Guide to Special Issues and Indexes of Periodicals* in 1962, 1976, and 1985. These guides detail the contents of consumer, trade, and technical periodicals. Within these guides, medical journals are given limited representation (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

Medline, the database maintained by the National Library of Medicine, and SciSearch, the science and technology database maintained by the Institute of Scientific Information, provide information which indicates what you can expect to retrieve from each database. Furthermore, guidelines for indexing collective titles are available in database indexing manuals (Charen 1976; Institute for Scientific Information 1993; National Library of Medicine 1992). However, indexing consistency studies show that variation exists even among experienced indexers who have access to these manuals (Funk, Reid, and McGoogan 1983; Leonard 1977). Many librarians, however, do not have access to the printed materials explaining indexing practices and policies (Conn and Poynard 1988; Wakeford and Roberts 1993).

Two 1990 studies by the McMaster University group compared searches conducted by end-users with those by librarians. The authors note that new passwords issued by the National Library of Medicine show that end-users constitute the fastest growing group of database users (McKibbon et al. 1990). In the studies, end-users received Medline training via the user-friendly software, Grateful Med. Both studies found that training increased the number of documents retrieved by end-users, but not the quality of retrieval. In the actual clinical setting, the time required in learning to search may exceed the amount of time a busy end-user can spare (Haynes et al. 1990; McKibbon et al. 1990). In addition, end-users usually do not have access to indexing manuals and other search aids. An earlier study found that when end-users have access to such materials, they still have difficulty formulating appropriate search strategies (Slingluff, Lev, and Eisan 1985).

Statement of Purpose

In the medical periodical literature, articles in series frequently do not have a common index term or subject heading which permits complete retrieval. Therefore, the first objective of this study is to measure the frequency with which articles in series can be retrieved by a series title.

One might also expect that a database indexed by humans (Medline), rather than a database maintained by automated indexing

(SciSearch), would better identify and index articles in series. Therefore, the second objective of this study is to measure Medline retrieval compared with SciSearch retrieval of individual articles in a series utilizing the series title.

SciSearch and Medline indexing manuals and search aids should provide the information necessary to predict retrieval of articles in series. The third objective of this study is to examine these materials and determine if expected retrieval of articles in series, based upon indexing guidelines and rules, differs from actual results.

Limitations

One year of selected periodicals was examined for articles in series. Series beginning or ending in 1993 were included in the study. Although all series are of interest to the medical community, only those which fit the typical format of a journal article were considered for inclusion in the present study. Therefore, directories, indexes, buyer's guides, convention reports, and product reviews were excluded.

The series included may appear as a permanent or temporary section of the periodical; however, all must have a unifying theme and an overall title. These criteria would exclude the main section of a periodical which contains original articles and/or brief reports representing many different themes. Criteria would exclude an entire issue devoted to a unifying theme but lacking an overall

title.

Also excluded are those sections of periodicals which contain articles representing specific publication types; for example, letters to the editor, case reports, editorials, review articles, or abstracts from other periodicals. A complete listing of excluded items can be found in appendix B.

CHAPTER II

LITERATURE REVIEW

Background Information

Historically, the word index was often synonymous with table, calendar, catalogue, inventory, register, summary, or syllabus. The concept is thought to have originated with the table of contents which accompanied medieval manuscripts. Only after the invention of the printing press was the arrangement to become alphabetical (Knight 1978). Famous English language eighteenth and nineteenth century indexes include Alexander Cruden's concordance of the bible, Samuel Johnson's *Dictionary of the English Language*, the twenty-second volume of the seventh edition of the *Encyclopaedia Britannica*, and Dr. John Shaw Billing's *Specimen Fasciculus of a Catalogue of the National Medical Library* (Blake 1986; Waserman 1972).

Cumulated Index Medicus and Medline

Specimen Fasciculus was expanded, refined, and ultimately published as *Cumulated Index Medicus* (CIM) in 1960 (Adams 1972). This international index is produced by the National Library of Medicine (NLM), a division of the U. S. Department of Health and Human Services located on the campus of the National Institutes of Health (Smith and Mehnert 1986).

CIM is published annually with monthly updates and contains citations to the biomedical journal literature. Indexed items include original journal articles and substantial letters, editorials, biographies, and obituaries. Items appearing in CIM are indexed by human beings. The indexers assign subject headings which describe the content of each item. Subject headings are selected from the NLM's controlled vocabulary, Medical Subject Headings or MeSH (National Library of Medicine 1992).

Science Citation Index and SciSearch

Science Citation Index (SCI) is produced commercially by the Institute of Scientific Information (ISI) in Philadelphia, Pennsylvania. It is published bimonthly with annual and five-year cumulations. Broader in scope than CIM, this international index contains citations to the science and technology literature as well as the biomedical literature. All items are indexed except advertisements, news notices, and most book reviews. SCI does not use a special nomenclature, classification system, subject headings, or thesauri. Instead, it indexes whatever natural language is expressed in titles, abstracts, and author-supplied keywords. In addition, it is not indexed by human beings; it utilizes automated indexing (Institute for Scientific Information 1993).

DIALOG Search Service

DIALOG is one of the few search services which provides access to both Medline and SciSearch. Although DIALOG's computers are located in California, access is provided through international telecommunications networks via local telephone numbers (DIALOG Information Services 1991b). SciSearch and Medline databases are structured and maintained by their respective producers, ISI and NLM, in very different ways. Consequently, there are subtle differences in the way in which each database is searched. DIALOG's contribution to all this diversity is to provide uniform commands for searching and extracting information. It is assumed that using a common search service, DIALOG, controls some of these variables.

Indexing the Journal Literature

General Information

Indexing the journal literature is the process of assigning terms to articles which represent the content of those articles. Terms are selected from a controlled vocabulary, as in Medline; or from natural language, as in SciSearch (Institute of Scientific Information 1993; National Library of Medicine 1992). The user may then retrieve articles by formulating a search request which utilizes those terms. Terms in a search request are connected by Boolean and proximity operators. The ensuing search produces a

collection of articles in which requested terms match assigned terms (Hersh and Greenes 1990).

Medline

Expert human indexers at the NLM assign terms to articles from the MeSH controlled vocabulary. MeSH consists of preferred terms, which may be assigned to documents, and entry terms (synonyms of preferred terms), which are not assigned. Preferred terms or MeSH headings may be further clarified by a set of subheadings. A MeSH heading-subheading combination generates a more specific search and, therefore, limits retrieval. MeSH headings exist in a hierarchical arrangement demonstrating relationships among terms. An 'explode' feature allows all MeSH terms under a more general heading to be included in the search. This feature generates a more inclusive search and increases retrieval (Hersh and Greenes 1990).

Although Medline is indexed by humans, a limited amount of computer assistance is given to the indexing process. In a 1987 article, Humphrey and Miller described computer-assisted indexing as it existed at the time of publication. Legitimate MeSH headings were verified and illegitimate terms were replaced with appropriate MeSH headings. Legitimate heading-subheading combinations were verified and illegitimate combinations were replaced with appropriate ones. In addition, computers assisted in identifying selected checktags (frequently used general headings such as HUMAN, MALE, FEMALE, ADULT) which always appear together; for example,

CHILD and HUMAN (Humphrey and Miller 1987).

The main focus of this article, however, was to introduce the MedIndEx system. This interactive knowledge-based system was created to expand the computer-assisted indexing of the medical literature so as to improve indexing consistency (Humphrey and Miller 1987). MedIndEx attempts to automate the indexing process as much as possible, but it should not be confused with the automated indexing utilized by SciSearch: the computerized processing of text into words and hyphenated word-phrases. Rather, the MedIndEx system prompts indexers to select appropriate MeSH headings and activates appropriate indexing rules (Humphrey 1988).

Several limitations of Medline's indexing and retrieval processes have been identified by Hersh and Greenes. First of all, end-users continue to have difficulty with the logical operators utilized in Boolean searching, even on the more user-friendly PaperChase and Grateful Med programs. Secondly, retrieval is not ranked according to its usefulness to the end-user or its relevance to the search request. A third limitation is that human indexing is expensive and time consuming. Lastly, some indexing inconsistency will no doubt remain in spite of the MedIndEx system (Hersh and Greenes 1990).

SciSearch

In contrast to Medline, SciSearch does not use a controlled vocabulary. Natural language, consisting of most title words,

author-supplied keywords, words in the abstract (since 1992), and title words and phrases from cited references (also since 1991), make up the SciSearch indexing system (Garfield and Sher 1993; Institute for Scientific Information 1993). The indexing of SciSearch is automated; it is not indexed by human beings. The automated indexing process employed by SciSearch should not be confused with that utilized by researchers developing new automatic indexing systems.

In SciSearch automated indexing removes frequently used but insignificant stop words (for example, AN, THE, and WHICH). These stop words may be full-stop, with complete removal, or semi-stop, with partial inclusion as co-terms. Retaining semi-stop words as co-terms facilitates "subheading" type combinations. Stop word removal was actually incorporated to eliminate millions of useless entries (Institute for Scientific Information 1993).

Automated indexing in SciSearch truncates words at eighteen characters for primary terms and eleven characters for co-terms (Institute for Scientific Information 1993). This occasionally has the same effect as the removal of suffixes with stemming algorithms (for example, -TION and -ING). Suffix and plural removal, or stemming, as utilized by automated indexing researchers, is employed to increase retrieval (Hersh and Greenes 1990).

Computerized word frequency statistics are utilized by SciSearch to identify words which are commonly associated with one another. These word combinations are then hyphenated as a phrase and eliminate an additional look-up when more than two

coordinations are necessary. This, of course, was designed explicitly for the printed format (Institute for Scientific Information 1993).

A new, unique feature of SciSearch's automated indexing, KeyWords Plus, resembles word frequency-based ranked retrieval. It is available for documents indexed after January 1991. This process extracts words and phrases from the titles of an article's cited references. These words are ranked and the top ten words or phrases serve to augment the title words, author-supplied keywords, and abstract words. KeyWords Plus words and phrases can then be used to search for additional relevant articles, thus increasing retrieval (Garfield and Sher 1993).

Although the automated indexing of SciSearch does not have the same disadvantages and limitations characteristic of the human indexing of Medline, authors Hersh and Greenes note several problems unique to SciSearch's indexing process. First of all, since there are no preferred terms which incorporate all possible synonyms, retrieving information in SciSearch requires the searcher to anticipate and include in the search statement all synonyms used to describe an entity. Furthermore, since there is no hierarchy of terms, all specific entities of a general concept (for example, all specific agents under ANTIBIOTICS) must be anticipated and included in the search statement. Finally, there are many word ambiguities (for example, LEAD the verb and LEAD the element) which produce irrelevant retrieval (Hersh and Greenes 1990).

Research and Development of Automatic Indexing Systems

Early research in automatic indexing focused on word frequencies. Luhn noted that words of high frequency (for example, AND or THE) and of low frequency were of little value in distinguishing relevant from irrelevant documents. The inverse document frequency formula considered not only the frequency of a word in an individual document, but also the frequency of the word in all documents in the collection. The weighting of a document's words, as determined by this formula, might be used to predict those words best employed in indexing the document (Hersh and Greenes 1990). Presently, there are three areas of research in automatic indexing: vector-based, probabilistic, and linguistic systems.

Vector-Based Systems. Salton, the proponent of the SMART vector-based system, uses word discrimination values to weight words in documents. In support of Luhn's theories, words of high and low frequency are poor discriminators; words of moderate frequency are better at distinguishing one document from another. A document vector is constructed from the controlled vocabulary of an article's abstract and a query vector is constructed from the natural language of the searcher. The vector cosine formula process matches query to documents and ranks them in descending order of similarity (Salton 1991).

Probabilistic Systems. The NLM's IRX Project is an example of a probabilistic system. It is used to evaluate experimental models which incorporate different weighting measures, stop lists, and stemming algorithms. All document and query words, except stop words, are weighted. When a query is put to the system, all documents which contain query words are retrieved. Weights of query words which appear in the documents are summed, and then documents are ranked according to their weights (Hersh and Greenes 1990; Hersh and Hickam 1992).

Linguistic Systems. Vector-based and probabilistic systems are sound mathematical models. They are successful at indexing and retrieving information. However, they are not as grounded in theory as linguistic systems. Linguistic systems attempt to index concepts rather than just words. They allow a greater variety of words to match concepts (Hersh and Greenes 1990).

In some linguistic systems, concepts and the relationships between concepts form a semantic network, where 'semantics' is the meaning of a word or phrase and 'network' is analogous to a classification system or hierarchy (Hersh and Greenes 1990). Vries and others created a thesaurus of neuroscience terms culled from the indexes of textbooks. Using semantic net expansion techniques, they successfully indexed a neuroscience subset of Medline articles (Vries et al. 1992).

It is generally agreed, however, that this system does not perform well in large domains characterized by diverse language

(Hersh and Greenes 1990). Hersh and others address this problem with concept-based indexing and retrieval using SAPHIRE. SAPHIRE processes natural language in both queries and documents. Concepts are identified and mapped to their canonical or preferred form using NLM's MetaThesaurus. The MetaThesaurus combines five controlled vocabularies: MeSH (National Library of Medicine), SNOMED (College of American Pathologists), DSM-III (American Psychiatry Association), ICD-9 (World Health Organization), and LCSH (Library of Congress). Retrieved documents, which match the user's query, are then ranked according to their relevance. Although somewhat inferior to Medline, SAPHIRE's indexing and retrieval shows promise (Hersh 1991; Hersh and Hickam 1992; Hersh and Hickam 1993).

In other linguistic systems, text is processed or parsed into syntactic (grammatical) categories (for example, noun phrases). The resulting syntactic categories are then used as index terms. These systems sometimes produce ambiguous and meaningless phrases (Hersh and Greenes 1990). Using a test collection of AIDS abstracts, Evans attempted to correct these problems by balancing the syntactic and semantic processing of text with CLARIT (Evans et al. 1991).

Indexing Journal Titles

General Information

Medline, the online counterpart of CIM indexes approximately

3,000 journals; SciSearch, the online counterpart of SCI, indexes approximately 4,000 journals (Marcaccio 1990). Both NLM and ISI rank the journals indexed in their respective databases. Medline indexes more thoroughly approximately 360 journals ranked as priority one. Journals of secondary priority, usually those outside the field of biomedicine, are selectively indexed (National Library of Medicine 1993). The labor intensive nature of Medline's human indexing process necessitates this prioritization.

It is not necessary, however, for SciSearch to prioritize journals. Their automated indexing process permits cover-to-cover indexing of all journals in a timely manner. Although not prioritized, SciSearch journals are ranked according to how frequently they are cited. This information is published annually in ISI's Citation Reports (Institute for Scientific Information 1991).

There are various other lists which rank medical periodicals according to their usefulness to health care professionals. These include Abridged Index Medicus and the Brandon-Hill List. Abridged Index Medicus is a list of the top 125 medical journals (National Library of Medicine 1993). Brandon and Hill provide a list of essential journals for the small hospital library which is published every other year in *Bulletin of the Medical Library Association*. Of the 143 journals on this list, some of the journals are marked as essential for initial purchase (Brandon et al. 1993).

A basic assumption of this research is that all articles in

series included in the study are indexed in both SciSearch and Medline databases and are considered important to health care professionals. Therefore, journals selected for study must be indexed by both SciSearch and Medline; categorized as priority one by NLM; included in Abridged Index Medicus; considered first purchase on the Brandon-Hill List; and, as an additional safeguard against indexing selectivity, ranked highest in ISI's Journal Citation Reports.

Medline

When accessing Medline via DIALOG, full and abbreviated journal titles are searchable with the JN=prefix; journal codes are searchable with the JC=prefix; and journal ISSN with the SN=prefix. Full journal titles are truncated at forty-six characters. A complete list of full and abbreviated journal titles along with their corresponding journal codes and ISSNs is given in the "Search Aids" section of Medline's *DIALOG Information Retrieval Service, White Sheets*. Journal titles must be searched as they appear on this list (DIALOG Information Services 1987).

SciSearch

When searching SciSearch via DIALOG, only full journal titles are searchable as complete phrases with the JN=prefix. Initial articles (a, an, and the) are dropped and titles are truncated at

forty-six characters. Exact punctuation and spacing must be used when searching directly, including hyphens and ampersands. If a journal contains logical operators (AND, OR, NOT) or the word FROM, the entire journal title must be enclosed in quotation marks. Journal titles may also be selected from the EXPAND command (DIALOG Information Services 1989).

Indexing Article Titles

General Information

There are many points of access to the literature in Medline and SciSearch databases via the DIALOG search service. In SciSearch one can use words from the abstract, the author-supplied keywords, KeyWord Plus, research fronts, and title words (DIALOG Information Services 1991a). In Medline one can use words from the abstract, subject headings, subheadings, check tags, special identifiers, a named person, and title words (DIALOG Information Services 1991a).

All such words are derived from either controlled or natural language vocabulary. Each has its own distinct limitations. In a controlled vocabulary, although subject headings group synonyms under a preferred term, they can be confining and may not respond to the natural evolution of language (National Library of Medicine 1992). In a natural language system, words from the abstract may be too diverse and contain insignificant and extraneous words (Kwok 1975). Title words express concepts at the whim of the individual

author with no control for synonyms and variants. Furthermore, title words may not accurately express the content of an article (Hodges 1983). Lastly, words from cited references display all of the previously mentioned difficulties (Salton and Zhang 1986).

In spite of their limitations, title words are frequently used for literature retrieval. They provide a simple, straight-forward method of retrieving medical information. In a 1961 study, title words alone adequately identified relevant articles pertaining to general research interests. However, abstracts as well as titles were required to answer specific questions (Resnick 1961). In a comparison of titles, abstracts, and full texts, Saracevic found that titles alone adequately identified non-relevant documents. However, when identifying relevant documents, abstracts were preferred over titles alone (Saracevic 1969).

To be effective title words must accurately represent the content of an article. Compared with the abstract, which contains numerous extraneous and insignificant words, titles are of limited length; consequently, they form a concise subset of those words found in the body of the article (Kwok 1975).

Although considered important, titles have not always been informative. In 1958, Luhn developed the key-word-in-context (KWIC) title index. This indexing method emphasized title word retrieval. After KWIC was introduced, it was thought that authors would begin to create more descriptive titles to insure that their published articles were discovered and used. Tocatlian studied the chemical literature from 1948 to 1968 in an attempt to support this

hypothesis. He compared substantive words in titles with the ten most substantive words in abstracts. Although his results were inconclusive, he was able to show that titles had become more meaningful and longer since 1948 (Tocatlian 1970).

When title words are used in combination with medical subject headings, retrieval is significantly enhanced. A 1991 study found that recall increased twenty percentage points when combination search strategies were used to retrieve double-blind trials in Medline (Gotzsche and Lange 1991). Retrieval effectiveness of titles, abstracts and subject headings in the COMPENDEX database found a combination of titles and abstracts came closest to 100 percent retrieval (Byrne 1975). A much older study, conducted in 1969, found a significant improvement in recall when title words and added keywords, rather than title words alone, were used to retrieve information from a collection of chemistry documents (Jahoda and Stursa 1969). In cases where indexers do not assign appropriate medical subject headings, title words not only enhance but frequently provide the only means of retrieval (Bernstein 1988; Poynard and Conn 1985).

SciSearch

In SciSearch all article titles and subtitles are indexed. Titles are examined by editors before the indexing process begins. Minor changes are made to improve title words as search terms. British spellings are standardized to American spellings; symbols

(for example, Greek letters) are spelled out fully; and place names, genus-species, and isotopes are hyphenated (Institute for Scientific Information 1993).

ISI uses word frequency statistics to identify words which are habitually hyphenated. These identified words are automatically indexed as fused words (for example, Viet-Nam is expressed as VIETNAM) and must be considered in the search strategy. On the other hand, words which consistently neighbor one another are hyphenated (for example, birth control is expressed as BIRTH-CONTROL) (Institute for Scientific Information 1993).

Medline

In Medline both British and American spellings may be used. Abbreviations, symbols, and acronyms as well as the complete word or phrase must be considered in the search strategy. Chemical formulas incorporate subscripts and superscripts as an alphanumeric string. Since words are not stripped of their suffixes, it is best to truncate in order to maximize retrieval (DIALOG Information Services 1987).

DIALOG

DIALOG has their own recommendations for article title searching in the SciSearch and Medline databases. All words in an article title can be searched except stop words which include: AN,

AND, BY, FOR, FROM, OF, THE, TO, and WITH. When stop words appear in an article title they must be replaced by the within-one-word proximity operator (1W). Words in an article title can be searched individually. When searched as multiple words, a proximity operator or the logical operator AND must be used. Finally, the searcher is cautioned about the use of acronyms as they may reflect different meanings in various disciplines (DIALOG Information Services 1989).

DIALOG treats hyphens and other punctuation as blanks, recommending the use of a proximity operator for words with internal punctuation (for example, insulin-like is searched as INSULIN(W)LIKE). For words not treated in a standardized manner (for example, nonlinear or non-linear), DIALOG suggests that the word be searched as a single word as well as a multiple word phrase with a proximity operator. An exception to this rule is the possessive apostrophe which is removed so that Reye's Syndrome is searched as REYES SYNDROME (DIALOG Information Services 1989).

Indexing Articles in Series

Guide to Special Issues and Indexes of Periodicals

The commitment of the Special Libraries Association to publish three editions of *Guide to Special Issues and Indexes of Periodicals* demonstrates the need for this type of information. More importantly, the guides show the diverse nature of special issues and sections published in periodicals, and suggest a format

for describing them. However, once computerized information retrieval became commonplace, the guides were discontinued.

The guides indexed consumer, trade, and technical periodicals. The first guide indexed 799 periodicals, the second 1,256, and the third 1,362. The second and third editions of the guide included both U. S. and Canadian periodicals and provided a separate classified listing of the periodicals indexed (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

All three guides included information about advertiser indexes, editorial indexes, and specials. Each edition defined specials somewhat differently. In the first edition specials included only annual sections, supplementary issues, and features. In the second edition the scope was broadened to include annual, semi-annual, or quarterly specials. In the third edition, the specials category was expanded to include directories, buyer's guides, convention issues, and statistical outlooks (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

The main portion of all three guides listed the indexed periodicals in alphabetical order. In the first edition, each periodical listed was given a code number, an indication of the periodical's frequency, and any associated organization. In the second and third editions additional periodical information included the address and price. The third edition alone provided the publisher and the periodical's online indexing and abstracting services (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

Each periodical entry included a brief description of those indexes and specials available. The availability of an advertiser index was stated. Information concerning the editorial index included frequency, form, and type (subject, title, author). In the first edition of the guide information about each special included title, brief annotation, and month released. In the second and third editions the price of the special, if separate from the periodical subscription, was noted. In the third edition alone the start date of the special was given (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

Each guide contained a detailed subject index which referred to the entry by code number. The first edition obtained the information included in the guide from questionnaires submitted to publishers. The second edition utilized various methods of gathering information: questionnaires, telephone conversations, actual examination of the periodical, editorial schedules, and other periodical directories (Devers, Katz, and Regan 1976; Katz, Madison, and Regan 1962; Uhlan 1985).

SciSearch

SCI documentation states that subtitles are indexed (Institute for Scientific Information 1993). Based upon this statement, all individual articles in a series are retrievable utilizing the series title if the individual article title is "attached" to the series title in a title-subtitle arrangement.

Medline

Statements in CIM documentation indicate that an "overall citation," with inclusive pagination, is created for those conference proceedings or abstracts, journal issues, or supplements devoted to a unifying theme (National Library of Medicine 1992). Based upon this statement, the issue or supplement of a journal, which has been devoted to proceedings or a unifying theme, is retrievable by series title. All other articles in series are apparently indexed by the following rather complex rule:

> If the title of the individual article when standing alone makes sense and reflects the substance of the over-all title, only the title of the individual article need be marked off [indexed]. If the title of the individual article when standing alone does not make sense and does not reflect the substance or meaning of the over-all title, the indexer must reproduce the over-all title . . . . Regardless of the nature of the source, follow the above procedure. This is based not on the identity of the individual paper in relation to the whole but on the sense of the title. . . . "Metabolism" is meaningless as a title when in reality the user should be oriented to "The biology of the Guinea Pig. Metabolism." Frequently, articles in a series are numbered. If the individual title can stand alone, DO NOT INCLUDE THE NUMBER in the marking [indexing]. If the individual title cannot stand alone, take [index] the over-all title, the number and the individual title. For example, with an over-all title reading, "Metabolism of the rabbit," "6. Copper metabolism of the rabbit" should be typed [indexed] as "Copper metabolism of the rabbit," not as "6. Copper metabolism of the rabbit" . . . . On the other hand, "Metabolism of the rabbit. 6. Copper" should be typed [indexed] as "Metabolism of the rabbit. 6. Copper" to give the reader the whole picture through both the over-all title and the item in the sequence (Charen 1976).

Operational Definitions of Series Studied

The aforementioned indexing rule reflects the complexity of

indexing articles in series. Defining just what constitutes articles in series is an equally difficult task. For the purposes of this study four types of series are identified: sections, supplements, issues, and sequences. A section is defined as a series of articles with and overall title appearing in an indefinite number of journal issues; a supplement is a series of articles with an overall title appearing in a supplemental issue of a journal; a issue is a series of articles with an overall title comprising the entire issue of a journal; and a sequence is a series of articles with an overall title appearing in a limited number of journal issues. Therefore, a section, by definition, represents a more permanent series than does the supplement, issue, or sequence.

Indexing Inconsistencies

It was earlier demonstrated that articles in series frequently do not share a common index term (see table 1). This unfortunate condition may not necessarily reflect indexing errors or inconsistencies but rather the diversity of the medical literature.

A similar situation occurs in the indexing of articles pertaining to syndromic entities not represented by a medical subject heading. *MEDLARS Indexing Manual* suggests that indexers assign approximately three subject headings to represent the dominant features discussed by the author of the article. In addition, indexers are directed to add the heading SYNDROME

(Jablonski 1992). Difficulties arise in selecting the dominant manifestations from among the many presented. Often the dominant features discussed simply reflect the author's or journal's special interests, rather than dominant manifestations of the non-MeSH syndrome (Jablonski 1992). Consequently, a group of these non-MeSH syndromes may often lack a common index term.

Indexing inconsistencies are not limited to syndromic entities. They occur in the medical literature in general, and may affect articles in series as well. Humphrey and Miller provide an excellent historical review of the various indexing consistency studies conducted in relation to the Medline database (Humphrey and Miller 1987).

The first indexing consistency study was conducted by Lancaster in 1968. Lancaster's study used three indexers to re-index sixteen articles. The second study, conducted by Leonard in 1975, used ten indexers and ten groups of articles (Leonard 1977). Authors Marcetich and Schuyler published a third study in 1981. They compared the indexing of fifty articles by four indexers who utilized a computer-assisted indexing aid (Associative Interactive Dictionary) with four indexers who did not use the aid. The indexing aid suggested medical subject headings based upon word frequencies in the abstracts of the articles (Doszkocs 1978). The fourth and last study was conducted by Funk and Reid. These authors used a much larger sample of 760 articles. The articles had been inadvertently indexed twice by NLM indexers, effectively eliminating the Hawthorne effect (Humphrey and Miller 1987).

All four studies used Hooper's formula for measuring indexing consistency (see figure 4). The studies clearly show that some inconsistency occurs when two different indexers index the same article. The greatest amount of consistency occurs in the assignment of checktags, geographic areas, and central concept headings. The greatest amount of inconsistency occurs in the assignment of main heading-subheading combinations (Funk, Reid, and McGoogan 1983).

$$CP\ (\%) = \frac{100\ A}{A + M + N}$$

A = number of terms in agreement
M = number of terms used by M, but not N
N = number of terms used by N, but not M

Figure 4.--Hooper's formula

Retrieval Difficulties

Indexing inconsistency and the absence of a series title, seems likely to produce incomplete retrieval of articles in series. In fact, Leonard's study firmly established a positive relationship between consistent indexing and effective retrieval (Funk, Reid, and McGoogan 1983; Humphrey and Miller 1987; Leonard 1977). A small but significant body of literature discusses difficulties encountered in retrieving information from the MEDLARS database. This literature does not directly involve articles in series. It is

concerned with retrieving articles which utilize randomized clinical trials (RCTs) in the research design.

In three studies published between 1985 and 1993, retrieval produced by a Medline search is compared with that produced by a manual search. Surprisingly, the results of the three studies demonstrate a greater yield with the manual rather than the Medline search. These unexpected results were then attributed to indexing inconsistencies of NLM's human indexers (Largaespada, Pistotti, and Bonati 1988; Poynard and Conn 1985; Silagy 1993). The conclusions asserted in the three studies produced numerous letters in response, pointing out various inadequacies in the search strategies utilized by the studies (Hewitt, Dickersin, and Chalmers 1988; Pinatsis 1988; Wakeford and Roberts 1993).

The origins of this controversy actually predate Medline with a paper written by Truelove and Wright in 1964. The authors present an excellent description of RCTs and explain how RCTs can be used in different pathological conditions of the gastrointestinal tract. Furthermore, they effectively demonstrate how RCTs fulfill all the requirements of the scientific method. The authors discover a scarcity of RCTs in the medical literature and predict that more RCTs would and should be conducted in the future (Truelove and Wright 1964).

A decade later Juhl and others decided to see if Truelove's prediction was correct. Utilizing Medline, the authors searched the gastroenterology literature and extracted those articles which met their criteria for RCTs. Although a comparison was not made between

manual and Medline search retrievals, both methods had to be utilized, as two years in the decade searched clearly predate Medline. The authors' findings substantiate Truelove's prediction. More importantly, however, the authors noted that less than 1 percent of the articles were incorrectly indexed as RCTs, resulting in "false positive hits." The authors then suspected that an unknown number of articles, which should have been indexed as RCTs, were not indexed as such. Unpublished data, confirmed their suspicions (Juhl, Christensen, and Tygstrup 1977).

Juhl's speculations stimulated Poynard and Conn to repeat the study approximately ten years later. The authors utilized similar RCT inclusion criteria and identical medical subject headings. This time, however, Medline retrieval of RCTs was compared with manual retrieval. Results favoring manual retrieval were found and a general dissatisfaction with Medline indexing of RCTs was voiced (Poynard and Conn 1985).

In 1988, Bernstein responded to the indexing inadequacies proposed by Poynard. Bernstein's study significantly modified Poynard's search strategy. Subheadings, utilized by Poynard, were eliminated by Bernstein because of Funk and Reid's earlier 1983 indexing inconsistency study (Funk, Reid, and McGoogan 1983). With increased retrieval in mind, Bernstein included corresponding anatomical terms for the broad disease entities originally utilized by Poynard. In addition, these terms were 'exploded' to capture the more specific anatomical and disease terms listed under them in the medical subject headings hierarchy. Bernstein's search strategy

also incorporated those changes in medical subject headings specific to the years searched. Lastly, the search strategy was adjusted to include text words in the titles and abstracts of articles (Bernstein 1988).

Introducing the concepts of recall and precision, Bernstein re-evaluated Poynard and Conn's results (see figures 5 and 6). In spite of incorporating the aforementioned heuristics to increase recall, Bernstein's search strategy did not recall significantly more articles than Poynard's search. However, Bernstein's search was significantly more precise than Poynard's. Although the author identified several non-indexing reasons for reduced retrieval of RCTs (for example, lack of searching skill, time constraints, and budget restrictions), she concludes that indexing inconsistencies prevent complete retrieval of RCTs (Bernstein 1988).

$$\text{Recall} = \frac{\text{number of relevant documents retrieved}}{\text{total number of relevant documents}}$$

Figure 5.--Recall formula

$$\text{Precision} = \frac{\text{number of relevant documents retrieved}}{\text{total number of documents retrieved}}$$

Figure 6.--Precision formula

The indexing inconsistency literature reveals a disturbing dissatisfaction with the indexing of the MEDLARS database. The studies demonstrate that NLM indexers, utilizing a controlled vocabulary, assign a significant number of "false positive and false negative" subject headings, resulting in insufficient relevant retrieval. When the appropriate subject headings are not assigned to RCT articles, retrieval is significantly improved with a title word search (Gotzsche and Lange 1991). Similarly, when a common subject heading is not assigned to each individual article in a series, retrieval is made possible with a series title word search. It is imperative, therefore, that the series title is indexed.

CHAPTER III

METHODOLOGY

Periodicals included in Abridged Index Medicus as priority one and the Brandon-Hill List as first purchase were identified in ISI's *Journal Citation Reports* (Institute for Scientific Information 1991), specifically "Journals Ranked by Times Cited in 1991." The top five general periodicals thus identified were included in the study (see figure 7).

1. **New England Journal of Medicine (NEJM)**
2. **Lancet**
3. **JAMA**
4. **Annals of Internal Medicine (Ann Intern Med)**
5. **American Journal of Medicine (Am J Med)**

Figure 7.--Periodicals Included in the Study

All issues of the selected periodicals published in 1993 were examined and articles in series were identified. Utilizing the database manager, Paradox (version 1.0), the information listed in figure 8 was maintained for each article in a series. The entire list represents a database record, each item on the list represents a database field, and indented items represent available selections for a field.

Information concerning each series was extracted from the completed article database and placed in a spreadsheet, utilizing Lotus123 (version 4.01). The information maintained for each series

is listed in figure 9. Unmarked items were supplied by the article database, items marked with an "*" were gleaned from perusing journal issues and indexes for 1992 and 1994, items marked with a "+" were provided by DIALOG searches in the Medline and SciSearch databases, and items marked with a '@' were derived from other information in the spreadsheet.

```
Article Title
Series Title
Series Title Type:
      Attached
      Unattached
Series Type:
      Issue
      Supplement
      Section
      Sequence
Journal Title
Journal Volume
Journal Issue
Journal Pagination
Journal Date
```

Figure 8.--Article Information

```
  Series Title
  Journal Title
  Series Title Type
  Series Type
  Series Frequency in 1993
* Series Occurrence in 1992-1994
  Number of Articles in the Series
+ Medline Retrieval
@ Relevant Medline Retrieval
@ Medline Recall
@ Medline Precision
+ SciSearch Retrieval
@ Relevant SciSearch Retrieval
@ SciSearch Recall
@ SciSearch Precision
```

Figure 9.--Series Information

BEST COPY AVAILABLE

Each series was searched by series title in both Medline and SciSearch utilizing the DIALOG search service. Since all of the series studied were published in the 1990s, only the more current file of each database was utilized. A search statement was created for each series title which excluded stop words, incorporated appropriate proximity operators, and designated only the title field. In addition, imbedded punctuation, truncation for plurals, and spelling variations were considered. Title search results were combined with the appropriate journal and date search results using the AND operator. All search results were printed and checked for correctness before retrieval was recorded.

Overall retrieval provided a measure of series title retrieval effectiveness. Retrieval was measured in Medline and SciSearch using formulas for recall and precision (see figures 5 and 6). The mean difference between Medline and SciSearch recall provided a measure of series title retrievability in a database indexed by humans compared with a database with automated indexing. Mean recall and precision for each type of series (issue, supplement, section, or sequence, each type of series title (attached or unattached), and series frequency and provided a measure of adherence to database documentation policies.

CHAPTER IV

ANALYSIS OF DATA

General Results

In the five journals examined, 134 series were identified from the 2,149 articles collected and stored in the article database. Table 2 presents data arranged by journal and series type. The largest concentration of series was found in JAMA (37 percent) and the second largest in Lancet (22 percent). In this study sections represent the most common type of series (73 percent), whereas issues and supplements the least common type (4.5 and 6.9 percent respectively).

Table 3 shows data arranged by series and title type. In general, series titles appear unattached more frequently (87 percent) than attached (13 percent). However, within the sequence category attached titles represent 77 percent of the total group.

Recall Results

Medline and SciSearch mean recall of series by title type is presented in table 4. The most obvious finding in this table is greater recall for series with attached titles (94 percent) than unattached titles (11 percent). In addition, mean recall of series with attached titles is the same for both databases. However, mean recall of series with unattached titles is nine percentage points

Table 2

Percentage of Total Series by Series Type and Journal

Series Type	Journal											
	Am J Med		Ann Intern Med		JAMA		Lancet		NEJM		All Journals	
	%		%		%		%		%		%	
Section	5.2	(7)	10.4	(14)	27.6	(37)	18.7	(25)	11.2	(15)	73.1	(98)
Supplement	4.5	(6)	1.5	(2)		(0)		(0)		(0)	6.9	(8)
Sequence	2.2	(3)	0.8	(1)	5.2	(7)	3.7	(5)	4.5	(6)	16.4	(22)
Issue		(0)		(0)	4.5	(6)		(0)		(0)	4.5	(6)
All Series	11.9	(16)	12.7	(17)	37.3	(50)	22.4	(30)	15.7	(21)		

Note: Figures in parentheses are base Ns for the adjacent percentages.
Total N=134.

Table 3

Percentage of Total Series by Series Type and Title Type

Title Type	Series Type									
	Section		Supplement		Sequence		Issue		All Sections	
	%		%		%		%		%	
Unattached	72	(97)	6	(8)	3	(5)	5	(6)	87	(116)
Attached	1	(1)		(0)	13	(17)		(0)	13	(18)
All Titles	73	(98)	6	(8)	16	(22)	5	(6)		

Note: Figures in parentheses are base Ns for the adjacent percentages.
Total N=134

Table 4

Mean Percent Recall and Precision of Articles in Series by Title Type

Title Type	Medline		SciSearch		Both Databases*	
	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)
Unattached (116)	16	20	7	17	11	19
Attached (18)	94	67	94	67	94	67
All Titles (134)	26	27	19	24	23	25

Note: Figures in parentheses are base Ns for the subsequent percents.
*Figures for both databases are based on 2N.

Table 5

Mean Percent Recall and Precision of Articles in Series by Series Type

Series Type	Medline		SciSearch		Both Databases*	
	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)
Section (98)	18	20	9	15	13	18
Supplement (8)	18	25	10	63	14	44
Sequence (22)	73	59	73	54	73	56
Issue (6)	4	17	0	0	2	8

Note: Figures in parentheses are base Ns for subsequent percents.
*Figures for both databases are based on 2N.

greater for Medline than for SciSearch.

Table 5 demonstrates mean percent recall of series by series type. This table shows mean recall of sections, supplements and issues four to nine percentage points greater for Medline than for SciSearch. However, mean recall of sequences is the same for both databases. This finding is consistent with mean recall of series with attached titles, and is not unexpected since 77 percent of sequences have attached titles.

Table 6 compares Medline and SciSearch percent retrieval of series with unattached titles at different proportions of recall: less than 50 percent, 50 to 99 percent, 100 percent, and any. In this table Medline retrieval exceeds that of SciSearch for complete recall by eight percentage points and for 50 to 99 percent recall by ten percentage points. However, any recall is essentially the same for both databases.

Table 6

Percentage of Series with Unattached Titles Retrieved by Percent Recall

Recall %	Medline Retrieval %	SciSearch Retrieval %
< 50%	9 (10)	15 (17)
50-99%	6 (7)	4 (5)
100%	10 (12)	2 (2)
Any	25 (29)	21 (24)

Note: Figures in parentheses are base Ns for the adjacent percents. N=116

The mean difference between Medline and SciSearch recall is

.07, and the standard deviation of this difference is .34. At a .05 level of significance, the slight positive difference in recall exhibited by Medline, is not statistically significant. Furthermore, the standard deviation of the difference indicates a wide variation from the mean difference between Medline and SciSearch recall.

Precision Results

In table 4, mean percent precision for both databases demonstrates greater values for attached titles (67 percent) than unattached titles (19 percent). When Medline and SciSearch are considered individually, mean percent precision of series with attached titles is the same. In general, Medline and SciSearch mean percent precision does not vary by any more than three percentage points in this table.

In table 5, mean percent precision for both databases demonstrates greater values for sequences (56 percent) and supplements (44 percent) than the other two categories of series types. When Medline and SciSearch are considered individually, more variation is evident in table 5 than in table 4. Medline mean percent precision is greater by five percentage points for sections and by seventeen percentage points for issues. SciSearch mean percent precision of supplements is greater by thirty-eight percentage points.

Correlation Between Recall and Precision

Correlation coefficients of recall and precision data were determined for both databases individually. The correlation coefficient of recall with respect to precision is .8 for Medline and .7 for SciSearch. Both databases exhibited a definite correlation between recall and precision. Furthermore, the relationship between recall and precision is positive rather than inverse: in general, as the values for recall increase, the values for precision also increase.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Retrieval Effectiveness

Although title searching is a useful tool in information retrieval, it is evident from this study that it is not an efficacious method of retrieving articles in series. Results demonstrate several reasons for this conclusion. First, the titles of most articles in a series are unattached (87 percent) to the series title (see table 3). Series with unattached titles are less likely to be indexed, according to database documentation (Charen 1976; Institute for Scientific Information 1992), and recall results given in table 4 tend to support this conclusion. Mean recall for series with unattached titles is only 11 percent, whereas recall for series with attached titles is 94 percent.

Second, when an unattached series title retrieves articles in a series, retrieval is typically incomplete. In table 6, when any amount of recall is considered, Medline retrieves only 25 percent of the series with unattached titles. Of that 25 percent only 10 percent represents complete recall. Similar values for SciSearch are 21 percent and 2 percent respectively.

Third, many series titles are general expressions and, as such, are not good candidates for discriminating retrieval. A comparison which demonstrates this point is *JAMA's* rather specific series title, "Clinical problem-solving," with a precision of 25

percent, and *Lancet's* more general series title, "Surgery," with a precision of only 4 percent. In addition, it was observed that titles of series with attached titles were more specific than titles of series with unattached titles. The results portrayed in table 4 serve to substantiate this observation. The precision of series with attached titles (67 percent) is considerably greater than the precision of series with unattached titles (19 percent).

Although generally not effective, there were occasions when recall by series title was substantial. The most obvious occurred when the series title was attached to the article title in title-subtitle configuration. Table 3 indicates 13 percent of series have attached titles. Table 4 demonstrates the substantial recall of series with attached titles: 94 percent for both Medline and SciSearch.

When recall of a series was 50 percent or greater (see table 6), it was observed that the series title and article title appeared in close proximity, although not actually attached. This was particularly evident in SciSearch's retrieval of *NEJM's* "Drug therapy," "Mechanisms of disease," Medical progress," and "Current concepts" (see table 7). These occasions suggest possible confusion in distinguishing when a series title is considered attached to an article title.

There were two large series with unattached titles in which Medline recall was 100 percent: *NEJM's* "Case records of the Massachusetts General Hospital" (fifty-two articles) and *JAMA's* "A piece of my mind" (thirty-three articles). In addition, there were

Table 7

Series with Unattached Titles and 50-99 Percent Recall

Database	Series Title	*Journal*	Recall %
SciSearch	Current Concepts	*NEJM*	65
	Drug Therapy	*NEJM*	87
	Mechanisms of Disease	*NEJM*	75
	Medical Progress	*NEJM*	67
	Tokyo Perspective	*Lancet*	50
Medline	Clinical Problem-Solving	*NEJM*	50
	From the Centers for Disease Control and Prevention	*JAMA*	84
	From the Congressional Office of Technology Assessment	*JAMA*	80
	From the National Institutes of Health	*JAMA*	67
	Images in Clinical Medicine	*NEJM*	55
	NIH Conference	*Ann Intern Med*	50
	Seminars in Medicine of the Beth Israel Hospital, Boston	*NEJM*	60

[a]Frequency refers to the percent of journal issues in which the series appears each year.

BEST COPY AVAILABLE

several occasions where Medline retrieval produced not only the series but the associated letters as well (*NEJM's* "Clinical problem-solving" and "Shattuck lecture"). These occasions represent a concerted effort on the part of NLM to index these sections according to instructions specific to each journal (Wright 1995).

It was also observed that some retrieval was entirely incidental. This occurred when the series title just happened to appear in the article title as well; for example, in *Lancet's* "Hypothesis" and "Surgery."

Recall and Precision Relationship

It is evident from results that an atypical relationship exists between recall and precision in this study. Instead of the usual inverse relationship (Fugmann 1985), there seems to be a positive relationship between recall and precision. Increases or decreases in recall are accompanied by corresponding increases or decreases in precision. This unexpected finding may be due to the small sample of data studied (only 134 series in five journals), or a lack of true randomization of the journals selected for study.

More likely, however, this finding is inherent to title word searching. First of all, search statements specified the title field which eliminated occurrences of words in other fields, such as the abstract or descriptors. Second, each word in the title was connected by proximity operators. The operators not only indicated proximity but a specific order as well. These heuristics produced

either 100 or 0 percent precision and recall in 79 percent of the Medline searches and 76 percent of the SciSearch searches (see table 8). This very specific title word searching may have skewed precision and recall towards a positive rather than inverse relationship.

Table 8

Percentage of Series with 0 or 100 Percent Recall and Precision

Database	0 Recall & Precision %	100% Recall & Precision %
Medline	66	13
SciSearch	69	7

N=134

Human vs Automated Indexing

The mean difference between Medline and SciSearch recall was the measure used to test the hypothesis of this study: a database which utilizes human indexing retrieves more articles in a series than a database which utilizes automated indexing. Since the slight positive difference in recall (.07) was not statistically significant, the hypothesis is rejected.

Support for this conclusion can be found in table 6. Medline demonstrates complete recall of 10 percent of series with unattached titles compared with 2 percent for SciSearch. However, SciSearch demonstrates incomplete recall of 19 percent of series with unattached titles compared with 15 percent for Medline. When

any recall is considered, both databases perform about the same.

These results suggest SciSearch retrieves an incomplete series more often than Medline, but Medline is more likely to retrieve a complete series. If the sample of journals investigated had been larger, perhaps the more numerous incomplete series retrieved by SciSearch would make a greater impact than the less numerous complete series retrieved by Medline. This conjecture would gain more support if higher-ranked journals, such as those investigated in this study, were given a more thorough deliberation by Medline than lesser-ranked journals.

Database Documentation

It is evident that recall of series with attached titles adheres to policies described in SciSearch database documentation (Institute for Scientific Information 1992). Recall of attached titles is 94 percent and represents the greatest recall for any group of data. Although titles and subtitles are often attached to one another by punctuation, they are frequently linked only by proximity or font. It is possible, then, that series title and article title combinations are mistaken for article title and subtitle. This may provide an explanation of why some of the substantial portions of the series with unattached titles listed in table 7 were retrieved by SciSearch.

Medline's 100 percent recall of *NEJM's* "Case records of the Massachusetts General Hospital" and *JAMA's* "A piece of my mind"

appears at first glance to deviate from indexing policies in database documentation. However, if the article titles in these two series are closely examined, it becomes obvious that they do not make sense when "standing alone" (Charen 1976). "Case records of the Massachusetts General Hospital" are simply assigned a case number which is the only title given. Typical article titles in the series, "A piece of my mind," are "Washing clothes," "In the footsteps," and "Babu."

The articles within supplements and special issues, which make up 10 percent of series, cannot be efficiently retrieved by series title. Supplement recall is 14 percent and issue recall is only 2 percent (see table 5). Generally, supplement results are in complete agreement with policies stated in Medline's documentation: the supplement itself is retrieved by the overall series title, but not the articles within the supplement (National Library of Medicine 1992).

A similar situation was discovered in two series found in *JAMA:* "From the Food and Drug Administration" and "From the National Institutes of Health." The entire section was retrieved by series title each time it appeared in an issue of *JAMA*, but not the articles within the section. However, the articles within these two sections were admittedly brief.

Recommendations

It is apparent that additional investigation is necessary

before the results found with the five journals studied can be generalized to all journals indexed by the two databases. Specifically, results from journals ranked lower by ISI should be compared with the results of the five journals included in this study.

Although documentation guides practice, confusion persists in distinguishing between a series title and an article title. Detailed indexing instructions are maintained on thirty-seven journals and brief notes on many of the remaining journals indexed at NLM. Specific instructions are given for the various sections included in the journal, including when to index both series title and article title (Wright 1995). Perhaps more predictable retrieval could be achieved if there was more cooperative effort on the part of database producers and journal publishers.

Two excellent series, "Facts, figures, & fallacies" and "Health and climate change," which appeared in *Lancet* cannot be retrieved by series title. "Facts, figures, & fallacies," like *British Medical Journal's* "ABCs of..." are marketed as separate publications as well as appearing in issues of their respective journals. This, however, is not a factor in NLM's treatment of these series (Wright 1995).

Special issues represent the least common type of series studied and the series least likely to be retrieved. Four of the most interesting issues examined appeared in *JAMA*. These four issues did not have an overall title and, therefore, were not included in the study. However, their content was definitely worthy

of mention. In each case an entire issue of *JAMA*, even letters to the editor and book reviews, was devoted to a topic of current importance: AIDS; human rights, war, and refugees; medical education; and genetics. Also, in 1992 two entire issues were devoted to violence. Unfortunately, issues such as these are retrieved only by happenchance.

There is a need to retrieve articles in series. This fact is aptly demonstrated by end-of-the-year indexes prepared by individual journals which include indexing of special sections and series. Although many series cannot be effectively retrieved by series title, it is frequently worth the effort.

APPENDIX A

Articles in Series Appearing in Figure 1

Arranged in Chronological Order

Facts, Figures, & Fallacies.

Jolley, Damien. 1993. The glitter of the t table. *Lancet* 342 (Jul): 27-29.

Victoria, Cesar G. 1993. What's the denominator? *Lancet* 342 (Jul): 97-99.

Grisso, Jeane Ann. 1993. Making comparisons. *Lancet* 342 (Jul): 157-160.

Carpenter, Lucy M. 1993. Is the study worth doing? *Lancet* 342 (Jul): 221-223.

Sitthi-Amorn, C., and V. Poshyachinda. 1993. Bias. *Lancet* 342 (Jul): 286-288.

Datta, Manjula. 1993. You cannot exclude the explanation you have not considered. *Lancet* 342 (Aug): 345-347.

Mertens, Thierry E. 1993. Estimating the effects of misclassification. *Lancet* 342 (Aug): 418-421.

Leon, D. A. 1993. Failed or misleading adjustment for confounding. *Lancet* 342 (Aug): 479-481.

Glynn, Judith R. 1993. A question of attribution. *Lancet* 342 (Aug): 530-532.

From the Centers for Disease Control and Prevention.

Centers for Disease Control. 1994. Distribution of STD clinic patients along a stages-of-behavioral-change continuum--selected sites, 1993. *JAMA* 271 (Jan): 2671-2672.

Centers for Disease Control. 1994. Update: mortality attributable to HIV infection among persons aged 25-44 years--United States, 1991 and 1992. *JAMA* 271 (Jan): 2672.

Centers for Disease Control. 1994. Assessment of street outreach for HIV prevention--selected sites. *JAMA* 271 (Jan): 2675.

Outpatient Parenteral Antibiotic Therapy. Management of Serious Infections. Part I: Medical, Socioeconomic, and Legal Issues

Tice, Alan D. 1993. Introduction. *Hospital Practice* 28 (Jun): 5.

Tice, Alan D. 1993. The team concept. *Hospital Practice* 28 (Jun): 6-10.

Brown, Richard B. 1993. Selecting the patient. *Hospital Practice* 28 (Jun): 11-15.

Craig, William A. 1993. Selecting the antibiotic. *Hospital Practice* 28 (Jun): 16-20.

Kravitz, Gary R. 1993. Advances in IV delivery," *Hospital Practice* 28 (Jun): 21-27.

Bradley, John S. 1993. Pediatric considerations. *Hospital Practice* 28 (Jun): 28-32.

Kunkel, Mark J. 1993. Quality assurance. *Hospital Practice* 28 (Jun): 33-38.

Milkovich, Gary. 1993. Costs and benefits. *Hospital Practice* 28 (Jun): 39-43.

Tierce, Jonothan C. 1993. Reimbursement. *Hospital Practice* 28 (Jun): 44-51.

Lawton, Stephan E. 1993. Legal issues. *Hospital Practice* 28 (Jun): 52-57.

Tice, Alan D. 1993. Discussion. *Hospital Practice* 28 (Jun): 58-64.

Users' Guides to the Medical Literature.

Guyatt, G. H., and D. Rennie. 1993. Users' guides to the medical Literature. *JAMA* 270 (Nov): 2096-2097.

Oxman, A. D., D. L. Sackett, and G. H. Guyatt. 1993. Users' guides to the medical literature I. How to get started. *JAMA* 270 (Nov): 2093-2095.

Guyatt, G. H., D. L. Sackett, and D. J. Cook. 1993. Users' guides to the medical literature II. How to use an article about therapy or prevention. A. Are the results of the study valid?

JAMA 270 (Dec): 2598-2601.

Guyatt, G. H., D. L. Sackett, and D. J. Cook. 1994. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? *JAMA* 271 (Jan): 59-63.

Jaeschke, R., G. Guyatt, and D. L. Sackett. 1994. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? *JAMA* 271 (Feb): 389-391.

Jaeschke, R., G. H. Guyatt, and D. L. Sackett. 1994. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* 271 (Mar): 703-707.

Levine, M., S. Walter, H. Lee, T. Haines, and others. 1994. Users' guides to the medical literature. IV. How to use an article about harm. *JAMA* 271 (May): 1615-1619.

Laupacis, A., G. Wells, W. S. Richardson, and P. Tugwell. 1994. Users' guides to the medical literature. V. How to use an article about prognosis. *JAMA* 272 (Jul): 234-237.

APPENDIX B

Sections Excluded From Study

Arranged in Alphabetical Order by Journal Title

American Journal of Medicine

Brief Clinical Observations
Case Reports
Clinical Studies
Correspondence
Editorials
Reviews

Annals of Internal Medicine

Articles
Brief Reports
Editorials
Letters
Reviews, Notes, and Listings

JAMA

Abstracts
Books, Journals, Software
Brief Reports
CME Forum
Corrections
Editorials
Instructions for Authors
Journal Club Reading
Letters
Obituary Listing/Obituaries
Original Contributions
Poetry and Medicine
Questions and Answers
Reference Directories
Resident Forum
The Cover

Lancet

Articles
Bookshelf
Correction
Diverticulum
Editorials
Film Review
In England Now
Letters to the Editor
News
News in Brief
Noticeboard
Obituary
People
Review Articles
Short Reports

New England Journal of Medicine

Books Received
Book Reviews
Brief Reports
Corrections
Correspondence
Editorials
Information for Authors
Notices
Original Articles
Review Articles

WORKS CITED

Adams, S. 1972. The way of the innovator: notes toward a prehistory of MEDLARS. *Bulletin of the Medical Library Association* 60 (Oct): 523-533.

Augustine, Patricia, Director, Medical Library. 1994. Interview by author, 15 November, Youngstown. Oral communication. Western Reserve Care System, Youngstown, OH.

Bernstein, Fran. 1988. The retrieval of randomized clinical trials in liver diseases form the medical literature: manual versus MEDLARS searches. *Controlled Clinical Trials* 9 (Mar): 23-31.

Blake, J. B. 1986. From Surgeon General's bookshelf to National Library of Medicine: a brief history. *Bulletin of the Medical Library Association* 74 (Oct): 318-324.

Brandon, Alfred N., Dorothy R. Hill, L. Gustave, and Janet W. Levy. 1993. Selected list of books and journals for the small medical library. *Bulletin of the Medical Library Association* 81 (Apr): 141-168.

Byrne Jerry R. 1975. Relative effectiveness of titles, abstracts, and subject headings for machine retrieval from the COMPENDEX services. *Journal of the American Society for Information Science* 26 (Jul-Aug): 223-229.

Charen, Thelma. 1976. *MEDLARS indexing manual. Part I. Bibliographic principles and descriptive indexing*. Bethesda: National Library of Medicine.

Conn, Harold O. and Thierry Poynard. 1988. Response to Hewitt et al: retrieval of randomized clinical trials in liver disease. *Controlled Clinical Trials* 9 (Mar): 88-90.

Devers, Charlotte M., Doris B. Katz, and Mary Margaret Regan. 1976. *Guide to special issues and indexes of periodicals*. 2d ed. New York: Special Libraries Association.

DIALOG Information Services. 1987. *DIALOG information retrieval service. White sheets*. Palo Alto: Dialog Information Services.

DIALOG Information Services. 1989. *DIALOG information retrieval service. White sheets*. Palo Alto: Dialog Information Services.

DIALOG Information Services. 1991a. *DIALOG information retrieval service. Blue sheets*. Palo Alto: Dialog Information Services.

DIALOG Information Services. 1991b. *DIALOG lab workbook*. Palo Alto: DIALOG Information Services.

Doszkocs, Tamas E. 1978. As associative interactive dictionary (AID) for online bibliographic searching. In *Proceedings of the ASIS annual meeting, 41st annual meeting, New York, NY, November 13-17, 1978, vol. 15*, by the American Society for Information Science. White Plains: Knowledge Industry Publications, 105-109.

Evans, David A., William R. Hersh, Ira A. Monarch, Robert G. Lefferts, and Steven K. Handerson. 1991. Automatic indexing of abstracts via natural-language processing using a simple thesaurus. *Medical Decision Making* 11 (Oct-Dec): S108-S115.

Fugmann, Robert. 1985. The five-axiom theory of indexing and information supply. *Journal of the American Society for Information Science* 36 (Mar): 116-129.

Funk, Mark E., Carolyn Anne Reid, and Leon S. McGoogan. 1983. Indexing consistency in MEDLINE. *Bulletin of the Medical Library Association* 71 (Apr): 176-183.

Gotzsche, Peter C. and Birgit Lange. 1991. Comparison of search strategies for recalling double-blind trials from MEDLINE. *Danish Medical Bulletin* 38 (Dec): 476-478.

Garfield, Eugene and Irving H. Sher. 1993. KeyWords Plus--algorithmic derivative indexing. *Journal of the American Society for Information Science* 44 (Jun): 298-299.

Gorman, M. and P. W. Winkler. 1988. *Anglo-American cataloging rules*. 2d ed. Chicago: American Library Association.

Haynes, R. Brian, K. Ann McKibbon, Cynthia J. Walker, Nancy Ryan, Dorothy Fitzgerald, and Michael F. Ramsden. 1990. Online access to MEDLINE in clinical settings. A study of use and usefulness. *Annals of Internal Medicine* 112 (Jan): 78-84.

Hersh, William R. 1991. Evaluation of Meta-1 for a concept-based approach to the automated indexing and retrieval of bibliographic and full-text databases. *Medical Decision Making* 11 (Oct-Dec): S120-S124.

Hersh, William R. and Robert A. Greenes. 1990. Information retrieval in medicine: state of the art. *Md Computing* 7 (May): 302-311.

Hersh, William R. and David H. Hickam. 1992. A comparison of retrieval effectiveness for three methods of indexing medical literature. *American Journal of the Medical Sciences* 303

(May): 292-300.

Hersh, William R. and David H. Hickam. 1993. A comparison of two methods for indexing and retrieval from a full-text medical database. *Medical Decision Making* 13 (Jul-Sep): 220-226.

Hewitt, Peg, Kay Dickersin, and Thomas C. Chalmers. 1988. More on MEDLINE searches. *Controlled Clinical Trials* 9 (Mar): 85-87.

Hodges, Pauline R. 1983. Keyword in title indexes. Effectiveness of retrieval in computer searches. *Special Libraries* 74 (Jan): 56-60.

Humphrey, Susanne M. 1988. *The MedIndEx system: research on interactive knowledge-based indexing of the medical literature*. Bethesda: Lister Hill National Center for Biomedical Communications. NTIS, PB88-243902.

Humphrey, Susanne M. and Nancy E. Miller. 1987. Knowledge-based indexing of the medical literature: the indexing aid project. *Journal of the American Society for Information Science* 38 (May): 184-196.

Institute for Scientific Information. 1991. *Science citation index. Journal citation reports*. Philadelphia: Institute for Scientific Information.

Institute for Scientific Information. 1993. *Science citation index. Guide and lists of source publications*. Philadelphia: Institute for Scientific Information.

Jablonski, Stanley. 1992. Syndrome--a changing concept. *Bulletin of the Medical Library Association* 80 (Oct): 323-327.

Jahoda, G. and Mary Lou Stursa. 1969. *Tests of indexes*. Tallahassee: Florida State University.

Juhl, Erik, Erik Christensen, and Niels Tygstrup. 1977. The epidemiology of the gastrointestinal randomized clinical trial. *New England Journal of Medicine* 296 (Jan): 20-22.

Katz, Doris B., Charlotte Madison, and Mary Margaret Regan. 1962. *Guide to special issues and indexes of periodicals*. New York: Special Libraries Association.

Knight, G. Norman. 1978. Book indexing in Great Britain: a brief history. In *Indexers on indexing: a selection of articles published in "The Indexer,"* ed. Leonard Montague Harrod, 9-13. New York: Bowker.

Kwok, K. L. 1975. The use of titles and cited titles as document representation for automatic classification. *Information*

Processing and Management 11 201-206.

Largaespada, Maria Jesus, Vanna Pistotti, and Maurizio Bonati. 1988. How accurate are bibliographic data bases? *Lancet* 1 (Mar): 538.

Leonard, L. E. 1977. Inter-indexer consistency studies 1954-1975: a review of the literature and summary of study results. In *Occasional papers, no. 131*, by the Graduate School of Library Science. Champaign: University of Illinois, 1-51.

Marcaccio, Kathleen Young, ed. 1990. *Computer-readable databases: a directory and data sourcebook*. 6th ed. Washington, D.C.: American Society for Information Science.

Maxwell, M. F. 1989. *Handbook for AACR2 1988 revision: explaining and illustrating the Anglo-American cataloging rules*. Chicago: American Library Association.

McKibbon, K. Ann, R. Brian Haynes, Cynthia J. Walker Dilks, Nancy C. Ryan, Lynda Baker, Tom Flemming, Dorothy Fitzgerald, and Michael F. Ramsden. 1990. How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. *Computers and Biomedical Research* 23 (Dec): 583-593.

National Library of Medicine. 1992. *Cumulated index medicus. Medical subject headings and journals indexed*. Chicago: American Medical Association.

National Library of Medicine. 1993. *List of journals indexed in "Index Medicus."* Washington DC: U. S. Department of Health, Education, and Welfare.

Pinatsis, A. 1988. Accuracy of bibliographic data bases. *Lancet* 1 (Apr): 942.

Poynard, T. and H. O. Conn. 1985. The retrieval of randomized clinical trials in liver disease from the medical literature. A comparison of MEDLARS and manual methods. *Controlled Clinical Trials* 6 (Dec): 271-279.

Resnick, A. 1961. Relative effectiveness of document titles and abstracts for determining relevance of documents. *Science* 134 (Oct): 1004-1006.

Rosenthal, Barbara, Director, Medical Library. 1994. Interview by author, 1 March, Youngstown. Oral communication. St. Elizabeth Hospital Medical Center, Youngstown, OH.

Salton, Gerard. 1991. Developments in automatic text retrieval. *Science* 253 (Aug): 974-980.

Salton, Gerard and Y. Zhang. 1986. Enhancement of text representations using related document titles. *Information Processing & Management* 22 (May): 385-394.

Saracevic, Tefko. 1969. Comparative effects of titles, abstracts and full texts on relevance judgments. *Proceedings of the American Society for Information Science* 6 (Oct): 239-299.

Silagy, Christopher. 1993. Developing a register of randomised controlled trials in primary care. *British Medical Journal* 306 (Apr): 897-900.

Slingluff, Deborah, Yvonne Lev, and Andrew Eisan. 1985. An end user search service in an academic health sciences library. *Medical Reference Services Quarterly* 4 (Spring): 11-21.

Smith, K. A. and R. B. Mehnert. 1986. The National Library of Medicine: from MEDLARS to the sesquicentennial and beyond. *Bulletin of the Medical Library Association* 74 (Oct): 325-332.

Tocatlian, Jacques J. 1970. Are titles of chemical papers becoming more informative? *Journal of the American Society for Information Science* 21 345-350.

Truelove, S. C. and R. Wright. 1964. The controlled therapeutic trial in gastroenterology. *American Journal of Digestive Diseases* 9 (Sep): 1-30.

Uhlan, Miriam. 1985. *Guide to special issues and indexes of periodicals*. 3rd ed. Washington, D.C.: Special Libraries Association.

Vries, John K., Bernadette Marshalek, Jennifer C. D'Abarno, Russell J. Yount, and Laura L. Dunner. 1992. An automated indexing system utilizing semantic net expansion. *Computers in Biomedical Research* 25 (Apr): 153-167.

Wakeford, R. and W. Roberts. 1993. Using Medline for comprehensive searches. *British Medical Journal* 306 (May): 1415-1415.

Waserman, M. J. 1972. Historical chronology and selected bibliography relating to the National Library of Medicine. *Bulletin of the Medical Library Association* 60 (Oct): 551-558.

Wright, Nancy, Director, Indexing Section. 1995. Interview by author, 28 March, Bethesda. Oral communication. National Library of Medicine, Bethesda, MD.